File 341
acc. 2228

A COMPARISON OF WH-CLEFTS AND *IT*-CLEFTS IN DISCOURSE

ELLEN F. PRINCE

University of Pennsylvania

WH-clefts and *it*-clefts have long been claimed to be interchangeable. However, an examination of naturally-occurring discourse demonstrates that this claim is incorrect. The presupposed part (WH-clause) of a WH-cleft represents information that the speaker can assume the hearer is thinking about. In one variety of *it*-cleft, the presupposed part (*that*/WH-clause) represents information which the speaker assumes the hearer knows or can deduce, but is not presumably thinking about. In another variety of *it*-cleft, the presupposed part represents information which the speaker takes to be a known fact, though definitely NOT known to the hearer.*

> 'There are situations where the speaker is constrained by a grammatical rule, and there are situations where he chooses according to his meaning ... ; but there are no situations in the system where "it makes no difference" which way you go ... This is just another way of saying that every contrast a language permits to survive is relevant, some time or other.' (Bolinger 1972b:71)

1. BACKGROUND. In the concluding lines of his monograph on *that*, Bolinger makes the strong claim quoted above. The following study is an attempt to provide evidence for the correctness of his claim with respect to two types of cleft sentences in English, the WH-cleft and the *it*-cleft; e.g.,

(1) a. John lost his keys.
b. What John lost was his keys.
c. It was his keys that John lost.

Corresponding to 1a, the simple or 'non-clefted' form, English has sentences like 1b, the WH-cleft, and 1c, the *it*-cleft.[1] They are all cognitively synonymous, each being true if and only if the other two are true; they therefore have the same objective information content. Thus a speaker who wishes to convey that information has, at least, the option of uttering any one of the sentences. Either the choice is random, and they are everywhere interchangeable—in which case Bolinger's claim is wrong—or else these sentences do different kinds of work and mean different things, at least some of the time.

There is a long literature in the field of transformational grammar on how cleft sentences are to be derived. For example, it was first assumed that each is derived

* I owe a great deal to Susumu Kuno, whose stimulating insights started me thinking about clefts in discourse. In addition, I should like to thank Catherine N. Ball, for sharing her data and ideas, as well as John Fought, Ivan Sag, Dwight Bolinger, Herbert Clark, Terry Langendoen, and Leigh Lisker for their helpful criticisms. A much abridged version of this paper was read at the LSA Winter Meeting, Chicago, 1977.

[1] This terminology is that of Hankamer 1974. *It*-clefts are also known as CLEFTS and WH-clefts as PSEUDO-CLEFTS. In the spirit of Higgins' 1971 syntactic characterization of the 'specificational pseudo-cleft', I am (perhaps arbitrarily) defining WH-clefts as sentences of the form *What* $S-C_i$ *is/was* C_i, where $S-C_i$ = Sentence minus Constituent$_i$. That is, I am excluding all clefts whose subject clause has a lexical head (*the one*, *the thing* ...) Likewise, *it*-clefts are here defined as *It is/was* C_i *which/who(m)/that/*$\emptyset$ $S-C_i$.

Austin Hall

directly from the unclefted version. Akmajian 1970 proposed that *it*-clefts be derived from WH-clefts, which Higgins 1971 argued are base-generated. Hankamer 1974 and Pinkham & Hankamer 1975 have given arguments for dual sources for each, one in which they are derived from the non-clefted structure, the other in which they are base-generated; Gundel 1977 argues that *it*-clefts are reduced forms of right-dislocated WH-clefts, and so on. It is not my intention here to argue for or against any proposed derivation, since it would seem that the final decision on their derivation must await the answer to the question of their equivalence—and, if they turn out to be non-equivalent, a clearer understanding than now exists of how the grammar is to account for the differences between them.

1.1. DIFFERENCES BETWEEN CLEFTS AND NON-CLEFTS. Although 1a–c all convey the same objective information in some crude sense, they differ in what has been called FOCUS and PRESUPPOSITION. Following Chomsky 1970 and others, we may say that 1b–c logically presuppose the proposition conveyed by 2:

(2) John lost something.

The criterion for presuppositionhood used here is one frequently used in transformational grammar: 'a sentence S presupposes a sentence S' just in case S logically implies S' and the negation of S, ~S, also logically implies S'' (Keenan 1971:45). To see that 1b–c both presuppose 2, but that 1a does not, we have only to consider the negations of 1a–c:

(3) a. John didn't lose his keys.
b. What John lost wasn't his keys.
c. It wasn't his keys that John lost.

Sentence 3a, but not 3b–c, may occur naturally and truthfully in a context like this:

(4) ... In fact, he's never lost anything in his life.

The fact that each of the clefted forms has *John lost something* as its presupposition and *his keys* as its focus has led many linguists of otherwise differing views to agree that WH-clefts and *it*-clefts are interchangeable. Thus Akmajian (149) says that 'they are synonymous, share the same presuppositions, answer the same questions, and in general they can be used interchangeably'; Bolinger (1972b:31) states that 'the distinction between them is ... "false"'; and Chafe (1976:37) writes that they 'evidently serve the same function'.

1.2. SYNTACTIC DIFFERENCES BETWEEN *IT*-CLEFTS AND WH-CLEFTS. Before comparing their discourse functions and domains, one must make note of and compare their syntactic domains, in order to recognize those differences in distribution that simply reflect arbitrary grammatical differences. As is well-known (see Akmajian 1970, Higgins 1971, Emonds 1970 etc.), the two types of cleft sentences differ in what categories of constituents may be focused. Both readily accept an NP (ex. 5), but an ADV (6) or a PP (7) commonly occur mainly (though not only) in *it*-clefts, while a VP (8) or an S (9) commonly occur in WH-clefts:

(5) a. '... it is THE GRAMMATICAL RELATIONS OF THE DEEP STRUCTURE that determine the semantic interpretation.' (Chomsky 1970:14)
b. 'What we have is simply A PRINCIPLE OF COOPERATION.' (Lasnik 1976:2)

(6) a. 'It was THEN that I became a young revolutionary.' (Hitler, transl. cited in Shirer 1960:13)
b. *When I became a young revolutionary was THEN.
(7) a. 'It is AGAINST PARDONING THESE that many protest.' (*PI*, 2/6/77, p. 6D)
b. *What/Where/How many protest is AGAINST PARDONING THESE.
(8) a. 'What that does is TEND TO ROB ERVIN AND THE GRAND JURY WITH YET A THIRD INVESTIGATING GROUP.' (Haldeman in *PT*, 201)
b. *It is TEND TO ROB ERVIN ... that that does.
(9) a. 'What you are saying is THAT THE PRESIDENT WAS INVOLVED.' (Haldeman in *PT*, 159)
b. *It is THAT THE PRESIDENT WAS INVOLVED that you are saying.[2]

Thus the only significant overlap concerns focused NP's. Even within the set of NP foci, the two constructions differ in that an *it*-cleft may focus an animate or inanimate NP with equal facility, whereas (in Modern English) a WH-cleft may focus only an inanimate NP. Compare 5 with the following:

(10) a. '... it was MAGRUDER that leaked it.' (Ehrlichman in *PT*, 221)
b. *Who/What leaked it was MAGRUDER.[3]

Of course, if a lexical head noun is added, animate NP's may also be focused; but

[2] Note, however, that sentences like (a) are acceptable; their focused constituents are adverbials:

(a) '... it is only TO DISCOVER THESE RELATIONSHIPS that the defendants ... are questioned so closely.' (Arendt 1951:434)

Furthermore, sentential subjects may be focused in an *it*-cleft, so long as they are potential surface subjects:

(b) It was that he would say such a thing that surprised me.
(c) *It was that he was sick that seemed to me.

I thank Ivan Sag for pointing out to me the possibility of sentences like (b). However, Terry Langendoen has informed me that, for him, any *that*-clause that functions as an NP can be clefted and, therefore, that he finds no difference in grammaticality between 9b, above, and (b).

In addition, as has been noted in the literature, one may focus certain predicate adjectives, as in

(d) It was green that she painted it.

[3] However, if the verb of the WH-clause is not selectionally restricted to human objects, then a focused human object NP is in fact acceptable:

(a) I don't need money—what I need is you.

Furthermore, non-referential NP's may always be focused:

(b) What he wants to be when he grows up is Pope.

Note also that the WH-cleft disambiguates the well-known ambiguity in sentences like:

(c) John wanted to marry a Norwegian.

Here *a Norwegian* may be either specific or non-specific. Only the latter may be focused in a formally-defined WH-cleft:

(d) What John wants to marry is a Norwegian. She $\left\{\begin{matrix}\text{*has}\\ \text{should have}\end{matrix}\right\}$ long hair.
(e) The one who John wants to marry is a Norwegian. She has long hair.

The situation with focused subject NP's is trickier: it seems, at least at first blush, that the focused human subject must be non-agentive. Compare the following:

then we no longer have a WH-cleft as formally defined:

(10) c. The one who leaked it was MAGRUDER.

1.3. WH-CLEFTS AND *IT*-CLEFTS IN DISCOURSE. When one studies actual cleft sentences in natural discourse, one is struck first by the differences in the average length of their focused and presupposed constituents. In the 37 WH-clefts with a focused NP that I have collected and tabulated, the presupposed string has one-third the average length of the focused string—or, put differently, the presupposed string constitutes approximately 25% of the corresponding whole non-clefted sentence. By contrast, in the 186 tabulated *it*-clefts in my corpus, the presupposed string is nearly twice as long as the focused string, or 66% of the corresponding non-clefted sentence. The actual figures are given in Table 1.

	ORAL				WRITTEN			
Type	No.	Focus	Presupp.	%	No.	Focus	Presupp.	%
WH	19	11.4	4.6	29%	18	19.3	5.3	22%
it + WH-	17	2.5	5.9	71%	41	5.7	11.2	66%
it + *th-*	71	2.8	3.3	55%	38	6.5	9.8	60%
it + ∅	19	2.1	3.3	61%	0	–	–	–

TABLE 1. Number of words in focus and presupposition and percentage of words in corresponding non-clefted sentence occurring in presupposition.

The figures under 'Focus' refer to the number of words in the constituent following the copula; the figures under 'Presupp.' refer to the number of words following *what* and preceding the copula in the WH-cleft, and following the WH-word (or *that* or ∅) in the *it*-cleft. The percentages refer to the percentage of words in the corresponding non-clefted sentence that occur in the presupposition of the cleft construction. When we analyse these figures, those for the WH-cleft seem intuitively correct if we follow the usual custom of equating logical presupposition with old/given/known information, and focus with new information.[4] That is, it seems natural to use as few words as possible to refer to something known, and to present new information in greater detail; this is in line with one of Grice's Maxims of Manner: 'Be brief (avoid unnecessary prolixity)' (1975:46). The figures pertaining to the *it*-clefts are, therefore, surprising, but not nearly so surprising as some of the actual examples:

(11) '... from a tissue of lies ... the notorious medieval piece known as the *Fetha Negest* was woven, thereby consecrating the theocratic and feudal character of the ruling classes. It is *this same myth* which IS PERPETUATED IN THE TWO SO-CALLED MODERN CONSTITUTIONS (THAT OF 1931 AND OF 1955), WHEREIN IT IS STATED THAT A MULTI-RELIGIOUS, MULTI-ETHNIC ETHIOPIA IS A CHRISTIAN FIEF CUT IN THE IMAGE OF THE RULING MINORITY'S OWN RELIGIOUS PREJUDICES AND CLASS INTERESTS.' (*Challenge*, p. 31)

It is obviously absurd to consider the presupposed clause (small capitals) as given or known information in the discourse, and/or the focused constituent (italics) as the new piece of information. In the same vein, it is clear that this sentence may not be replaced by a WH-cleft in this discourse, preserving coherence (however that is to be defined). Thus we see that, even limiting ourselves to focused NP's, discourse-related differences exist between the two types of clefts. We shall now examine each type separately, in the hope of discovering its particular functions, and the discourse factors relevant to it.

2. WH-CLEFTS. If we compare the logico-semantic notion of presupposition with the discourse notion of known/old/given information, we find that, in the case of the WH-cleft, they seem to match quite closely. In some instances, the information is in fact given in the preceding linguistic context:

(12) 'There is no question what they$_i$ are after. WHAT THE COMMITTEE$_i$ IS AFTER is somebody at the White House. They$_i$ would like to get Haldeman or Colson, Ehrlichman.' (Nixon in *PT*, 64)

However, such repetition is quite rare; the vast majority of cases refer to information which is known or knowable from implications/implicatures:

(13) a. 'He is reported ... not to be as desperate today as he was yesterday but to still be on the brink, or at least shaky. WHAT'S MADE HIM SHAKY is that he's seen McCord bouncing out there and probably walking out scot free.' (Haldeman in *PT*, 186)

b. 'Himself a religious Jew, Prof. Flusser says that Carter's piety is not the problem. "WHAT I'M WORRIED ABOUT", he declares, "is that the influence of Carter's advisers may eclipse the influence of the Divine Spirit."' (*JE*, 9/10/76, p. 33)

c. 'Nikki Caine, 19, doesn't want to be a movie star. WHAT SHE HOPES TO DO is be a star on the horse-show circuit.' (*Today*, 10/10/76, p. 44)

In 13a, we are informed that 'he' is shaky, and we need infer only that something has made him so. In 13b, the situation is slightly more complex: the first sentence informs us indirectly that Flusser (and perhaps the writer) thinks that there is a problem (compare with *not a problem*), and it informs us directly that Carter's piety is not it. The only inference or BRIDGE (Haviland & Clark 1974) that the reader must construct is that, if one knows of a problem, then one worries about it.

The notion of bridge—the inference one must make to get from the new information back to what is known—is particularly useful for 13c: the first sentence could conceivably (though not very coöperatively) be followed by something like *In fact, she doesn't want to be anything at all*. This shows that one cannot KNOW from the first sentence that N.C. wants to do or be something. When one hears the cleft, however, one simply constructs an inferential bridge—N.C. wants to do something—which is actually quite compatible with our knowledge of the world and of 19-year-olds.

(f) John made me laugh.

(g) What made me laugh was John.

In (f), *John* may be read as agentive or non-agentive (instrumental?); but only the non-agentive reading is possible in (g). This is in contrast with the situation that obtains in cleft sentences with lexical heads, where agentive readings are possible:

(h) The one that made me laugh was John.

This is an area requiring further study.

[4] For example, 'The stuff inside the WH clause must be presupposed (given from context) in contrast to the focus, which is new information' (Hankamer, 223; see also Quirk & Greenbaum 1973:414 ff., Hooper & Thompson 1973:495, and Gundel 1977:533).

Austin Hall

The relationship between WH-clefts and known information leads Hankamer (223) to suggest that the former cannot occur discourse-initially. However, consider the following (## marks what follows as discourse-initial):

(14) ##'What we have set as our goal is the grammatical capacity of children—a part of their linguistic competence.' (Klima & Bellugi-Klima 1969:448)

This is the opening of their article on syntactic regularities in the speech of children. Obviously, nothing in the title lets you know specifically that the authors have set something as their goal. Nevertheless, the bridge is constructible, not from the linguistic context but from the situation or genre: the reader of a linguistics article ASSUMES that its author had a goal in writing the article.

So far, we have looked at real data and have tried to account for their acceptability. However, consider the following invented discourses:

(15) a. ##*Hi! What my name is is Ellen.
b. ##*Hi! What I've heard about is your work.
c. ##*Hi! What you used to do was go to school with my brother.

If 15a–c were acceptable (as discourse openers), we would have no trouble explaining why: the hearer, seeing that I am a person, could infer that I have a name (15a) and even that I've heard about something (15b), and s/he could most certainly build a bridge to the effect that s/he used to do something (15c). This is a most unfortunate situation, as these discourses are clearly unacceptable. Furthermore, note that the corresponding non-clefted forms are fine:

(16) a. ##Hi! My name is Ellen.
b. ##Hi! I've heard about your work.
c. ##Hi! You used to go to school with my brother.

Clearly, there are constraints on bridge-building in the case of WH-clefts. If we reconsider discourses like 14 and compare them with (non-)discourses like 15, we note a difference: WITH ALL DUE MODESTY, THE WRITERS OF 14 CAN ASSUME THAT THE READER WILL HAVE IN MIND, AT THE MOMENT OF READING THE OPENING, THE FACT THAT THEY HAVE A CERTAIN GOAL—i.e., the bridge is not only constructible, but can politely (though perhaps incorrectly) be assumed to be already present in the reader's mind or CONSCIOUSNESS, following Chafe 1974. By contrast, in 15a, it would be highly presumptuous (though possibly correct) for me, the speaker, to assume that my hearer is actually thinking that I have a name; and it would be absurd to assume that s/he has in mind the bridges necessary in 15b–c. Thus we may posit the following discourse condition on WH-clefts:

(17) Discourse condition on WH-clefts: A WH-cleft will not occur coherently in a discourse if the material inside the (subject) WH-clause does not represent material which the coöperative speaker can assume to be appropriately in the hearer's consciousness at the time of hearing the utterance.

2.1. WHAT CAN BE ASSUMED to be appropriately in the hearer's consciousness depends, in large part, on the speech situation. A college professor may felicitously begin a lecture (or a course) with 18a, but not usually with 18b:

(18) a. ##What we're going to look at today (this term) is ...
b. ##*What one of my colleagues said this morning was ...

Likewise, students, showing up for appointments with an adviser, may start with 19 if it was they who requested the meeting, but not usually if it was the adviser:

(19) ##Well, what I wanted to talk to you about was ...

In a different context, consider Smith, who rings the doorbell of neighbor Jones, with whom he has had hardly any contact; 20a would be acceptable, but not 20b:

(20) a. ##Hi! I wonder if I might borrow a cup of sugar.
b. ##*Hi! What I wonder if I might borrow is a cup of sugar.

But if the situation is such that Smith habitually borrows things from Jones, then it is appropriate for Smith to assume that his act of ringing Jones' bell is sufficient to make Jones think that Smith wants to borrow something, in which case 21 would be acceptable:

(21) ##Hi! What I wonder if I might borrow this time is a cup of sugar.

The important point is that the particular piece of information or bridge or ANTECEDENT (Clark & Haviland 1977) does not actually have to BE in the hearer's consciousness; in fact, the speaker may at times be fairly certain that it is not. It simply has to be appropriate to the situation, and hence coöperatively 'assumable' as being there. One might think that the set of possible antecedents is totally uncharacterizable and unpredictable; but my data show that speakers and speech situations are not as idiosyncratic as they might be. The following are the major types of information units which may coherently serve as antecedents to WH-clefts.

2.11. EXPLICIT INFORMATION, MORE OR LESS. This includes information given explicitly (or nearly so) in the preceding linguistic context or in the accompanying non-linguistic context, as in 12 above and the following:

(22) a. '... in Haviland and Clark (in press) they say that ... Certainly WHAT THEY ARE TALKING ABOUT IS ...' (Chafe 1976:42)
b. 'You see, WHAT I AM DOING, John, is putting you in the same situation as Pres. Eisenhower put me in with Adams.' (Nixon in *PT*, 278)

2.12. IMPLICIT INFORMATION. Since actual phrases are rarely repeated, the line between explicit and implicit information is often fuzzy. The hearer has a bit more work to do in discourses like 23 than in ones like 13a or 22 above:

(23) a. 'At first contact he developed a furious hatred for the party of the Social Democrats. "WHAT MOST REPELLED ME", he says, "was its hostile attitude toward the struggle for ..."' (Shirer, 22)
b. 'When talk goes on of "the national purpose", "the national life", "the mainstream of the nation", WHAT IS BEING ADVANCED is the erection of these bourgeois dictatorships.' (*Challenge*, p. 8)
c. [speaking of the impossibility of a 20-hour work-week] 'The intellectuals, they always say there are potential Lord Byrons, Walt Whitmans, Roosevelts, Picassos working in construction or steel mills or factories. But I don't think they believe it. I think WHAT

THEY'RE AFRAID OF is the potential Hitlers and Stalins that are there too.'[8] (Steelworker in Terkel 1974:4)

2.13. CONTRAST. The following discourses are similar to those of §2.12 in that their antecedents are related to the preceding linguistic context, but here the relation is one of contrast. The contrast may be one of negative/affirmative, antonymy, positive/comparative/superlative, time etc. Consider 13b–c above and the following, where the small capitals mark the items contrasted:

(24) a. 'Precisely how pseudo-clefts are formed NEED NOT CONCERN US ... WHAT IS RELEVANT is that in all the cases examined above—and in fact in most pseudo-clefts—the constituent following *be* is an NP.' (Akmajian & Heny 1975:316)

b. 'Our position is a DYNAMIC one. It will be more and more refined as conditions change in the course of the struggle. WHAT IS CONSTANT is our commitment to a revolutionary emancipation of Ethiopia.' (*Challenge*, p. 1)

c. 'The fact that ... pre-eminence of some groups and regions over others shifted frequently is WELL KNOWN ... WHAT IS LESS KNOWN, OR RATHER NOT ADMITTED BY SOME WHO PREFER NOT TO LOOK AT THE STARING PRESENCE OF REALITY, is the other fact that ...' (*Challenge*, p. 25)

d. 'WHEN WE FIRST STARTED TO RESEARCH WOMEN'S PART IN HISTORY, we concentrated on the superstars ... WHAT HAS BECOME MORE IMPORTANT TO US NOW, THOUGH, is how the average woman lived ...' (*PB*, 7/23/76, p. 27)

Until now, the antecedents have been more or less (re)constructible on semantic grounds, but pragmatics is far more important in the categories below.

2.14. METALINGUISTIC ANTECEDENTS. This group of WH-clefts, like those above, has antecedents which can be (re)constructed from the preceding context; but now the hearer is assumed to be entertaining certain METALINGUISTIC notions. In particular, if A is/has been speaking or is reported to have spoken, then anyone involved in the speech situation (including A) can coöperatively assume that everyone else so involved has in mind that A means/meant something. Thus we find many WH-clefts in contexts like the following:

(25) a. Nixon: '... There is something to be said for not maybe this complete answer to this fellow, but maybe just a statement to me. My versions are these: bing, bing, bing. That is a possibility.'
Dean: 'Uh huh.'
N.: 'WHAT I MEAN is we need something to answer somebody.' (*PT*, 95)

b. Dean: '... There have been a lot of people having to pull oars and not everybody pulls them all the same time, the same way, because they develop self-interests.'

[8] The speaker of 23c may have felt that he was presuming too much on the part of his hearer (an 'intellectual'), or that he had glossed over an important point, for he continued the discourse with: 'The people in power fear the leisure man.'

Haldeman (to Nixon): 'WHAT JOHN IS SAYING, everybody smiles at Dean and says, "Well, you better get something done about it."'
D.: 'That's right.' (*PT*, 129)

c. Nixon: '... I would hope to keep the Grand Jury from leaking, but—'
Petersen: 'Well, you know, I don't want to go too far on it either, because I don't want to get into a diversionary battle with Anderson.'
N.: 'Hell, no. I wouldn't pay that much attention to it. I agree, I agree. WHAT I MEANT is just do the best to control it.' (*PT*, 637)

It is significant that there is often no complementizer *that* in the focused clause of such sentences. In the same vein, the focused clause often displays certain other features normally associated only with main clauses, e.g. the imperative in 25c and certain of the so-called root transformations. Since such WH-clefts are often used not simply for clarifying previous assertions, but also for remaking them, the presence of such main-clause phenomena lends support to a functional explanation of their existence, in the manner of Hooper & Thompson 1973 and Green 1976; and the lack of *that* (also construable as a main-clause phenomenon) is consistent with Bolinger's claim (1972b) that *that* is really anaphoric.

Another, less common, type of metalinguistic WH-cleft is exemplified by the following:

(26) 'The bourgeois status of the right to self-determination is not affected by the fact that the bourgeoisie was historically able to marshal the help of the peasantry and proletariat ... WHAT MAKES THE DESIGNATION "BOURGEOIS" APPROPRIATE TO THIS CLAIM TO SELF-DETERMINATION is that ...' (*Challenge*, p. 11)

If a coöperative speaker labels something as 'bourgeois', e.g., s/he may then assume that the hearer is thinking that—according to the speaker, at least—such a label is appropriate.

The four groups of WH-clefts presented above all have their antecedents contained in or at least partly (re)constructible from their preceding contexts. In the other types of WH-clefts that I have found, the antecedents lie in the norms of the speech situation rather than in the texts themselves.

2.15. *THINK* ... In ordinary, non-ritualized discourse, it seems that the speaker's relevant thoughts, observations, opinions, reactions etc. (often negative) are taken to be the constant appropriate concern of the hearer, and may therefore nearly always be assumed to be in the hearer's consciousness; e.g.,

(27) a. Nixon: '... I knew there was something going on, but I didn't know it was a Hunt.'
Dean: 'WHAT REALLY TROUBLES ME is: one, will this thing not break some day and the whole thing—domino situation—everything starts crumbling, fingers will be pointing ...' (*PT*, 115)

b. Nixon: '... Dean was in the news, I guess.'
Ehrlichman: 'Now wait a minute. All right, not as much as Magruder. Magruder, too. WHAT IS SHOCKING TO ME is his blowing off against the one fair guy you wouldn't think he would cut up, against Haldeman.' (*PT*, 209–10)

Austin Hall

As might be expected, this type of WH-cleft is most common in conversational discourse. However, it is clear that not all sentences beginning with *I think* etc. may be WH-clefted:

(28) a. A: Wasn't that incredible when Mary called the boss a pig?
B: Yeah, it really shocked me that she called him that.
Yeah, that she called him that really shocked me.
*Yeah, what really shocked me was that she called him that.

b. A: Is Sam here yet?
B: No, I don't believe he's arriving until Tuesday.
*No, what I don't believe is that he's arriving until Tuesday.

c. A: Hello, Operator. I'm trying to dial BU7-1151. Could you please check it for me?
B: I think the exchange is overloaded. Hold on while I check it.
*What I think is that the exchange is overloaded ...

The WH-cleft in 28a is unacceptable because it marks the information 'something really shocked me' as present in the hearer's consciousness—in contrast to the information 'she called him that', which in fact is more immediately present in his/her consciousness. (Note that an inverted WH-cleft would, however, be acceptable here.) While it may be too strong to claim that the focus in a WH-cleft must present absolutely new information, it clearly must not be assumed to be MORE prominent in the hearer's consciousness than the information in the WH-clause.

In 28b, *I don't believe* is used metaphorically as a hedge (see Prince 1976) and, therefore, is not literally about the speaker's beliefs. Note that such hedges may, however, be properly included in a WH-clause:

(29) I would do a lot of crazy things, but WHAT I DON'T THINK I WOULD DO is buy a car until I could pay for its upkeep.
...*but WHAT I DON'T THINK is that I would buy a car until I could pay for its upkeep.

In the acceptable version of 29, the antecedent is a hedged 'I wouldn't do something', (re)constructed by means of contrast from the preceding clause. In the unacceptable version, the WH-clause is made up entirely of the hedge, which is too weak informationally to represent something that the hearer is supposed to be thinking about. Note that the scope of the negation does not seem to be a relevant factor.

In 28c, I would attribute the strangeness of the WH-cleft to the genre or speech situation. While one is supposed to be actively concerned with the relevant thoughts of one's co-participants in normal conversation, this is not necessarily so in highly ritualized service encounters (see Merritt 1976). Thus a telephone operator may inform me that s/he thinks something, but s/he cannot coöperatively assume that the fact that s/he thinks something is already present in my consciousness.

2.16. *HAPPEN*. Note the following examples:

(30) a. Haldeman: 'He said, yes, he thinks John Dean did lie to the FBI when he said he wasn't sure whether Howard Hunt had an office in the White House.'

Dean: 'I said I had to check it out. WHAT HAPPENED is that the agent asked if he could see the office ...' (*PT*, 153)

b. '... with a number of intransitive Iroquois verbs there appears to be an agent prefix agreeing with a patient noun. Such is the case with the first person agent prefix *-k-* in *oʔkáʔsēʔt* ... WHAT SEEMS TO BE HAPPENING is that agenthood and patienthood have become partially confused ...' (Chafe 1976:47)

The antecedents 'something happened/seems to be happening' are not retrievable from anything in the preceding contexts; yet the discourses are coherent. There seems to be a pragmatic principle that says that events keep occurring—and that in our culture, at least, they are our proper and constant concern. *Happen*-clauses are extremely common in the oral data, and not at all infrequent in the written. Their distribution in different types of discourse is wider than that of *think*-clauses, indicating that it is more generally appropriate for a speaker to assume that his/her co-participants are consciously concerned with past, present, and future events than with his/her thoughts. Compare, e.g., 28c with the following:

(31) A: Hello, operator. I'm trying to dial BU7-1151. Could you please check it for me?
B: [checks] WHAT HAPPENED is that the whole exchange got overloaded. Wait a few hours and try again.

Like the *think* and metalinguistic clefts, *happen*-clefts frequently introduce main assertions, and thus often occur without the *that*-complementizer:

(32) Nixon: '... and I think WHAT HAPPENED is he ran that string out and finally just gave up because that wouldn't work and ...' (*PT*, 407–8)

2.2. CONSCIOUSNESS AND GIVEN INFORMATION. The notion of consciousness was, to my knowledge, introduced in the literature by Chafe (1974:111–12) where it is identified with GIVEN INFORMATION:

> 'Language ... is used primarily to increase the amount of knowledge that is shared by separate minds ... What a speaker shares with his addressee must be part of what is in the speaker's consciousness at the time ... The speaker knows he is not introducing material from his own consciousness into an empty vessel, but that his task is to introduce new things into a consciousness (the addressee's) which already has some content ... The speaker must make assumptions as to what the addressee is conscious of, and transmit his own material accordingly. Virtually every sentence ... is a mixture of what, following Halliday 1967, I will call GIVEN material, which the speaker assumes is already in the addressee's consciousness, and NEW material, which he assumes is not.'

Chafe goes on to state that given and new information are treated differently, i.e. are formally distinguishable; and he gives low pitch and pronominalization as examples of markers of givenness. Thus Chafe distinguishes the given, what is assumed to be in the hearer's consciousness, from the new, its complement.

Halliday 1967 approaches the notion of given vs. new from the opposite direction, starting with the notion of INFORMATION FOCUS: 'that which is represented by the speaker as being new, textually (and situationally) non-derivable information' (p. 205). Information focus may be marked (e.g. by special stress) or unmarked (e.g.

by 'normal' stress), and given information is the complement of the MARKED focus only (208):

> 'A distinction may ... be made between unmarked focus, realized as the location of the tonic on the final accented lexical item, which assigns the function "new" to the constituent in question, but does not specify the status of the remainder, and marked focus, realized as any other location of the tonic, which assigns the function "new" to the focal constituent and that of "given" to the rest of the information unit.'

Thus Chafe and Halliday define GIVEN in entirely different ways. Chafe takes the speaker's assumptions as primitive, and directly relates givenness to certain of those assumptions; Halliday, if I understand him correctly, takes intonation as primitive, and indirectly relates givenness to a certain intonation contour.[6]

With respect to the problem at hand, it is Chafe's notion of givenness—THAT WHICH THE SPEAKER ASSUMES TO BE IN THE HEARER'S CONSCIOUSNESS—that is pertinent to WH-clefts. The one change I would make is from 'assumes' to 'may coöperatively assume,' since the speaker need not actually think that the information is in the hearer's mind—only that it would be appropriate if it were. It is only to such information that the coöperative utterer of a WH-cleft may require bridge-building on the part of the hearer.

3. *IT*-CLEFTS. In contrast to the WH-clefts, the information represented in *it*-cleft *that*-clauses does NOT have to be assumed to be in the hearer's mind, although of course it may. Consider the following:

(33) a. 'If I see a train crossing, I keep going. IT'S A GAME YOU'RE PLAYING. Watch the stop light ...' (Utility man in Terkel, 233)
b. 'I've been bit once already by a German shepherd. It was really scary. IT WAS AN OUTSIDE METER THE WOMAN HAD. I read the gas meter and was walking back out ...' (Meter reader in Terkel, 366)
c. E.P.: 'Mmm ...' [eating pieces of fudge]
K.M.: 'Aren't those good? IT WAS ONLY SHEER WILL POWER THAT KEPT ME FROM EATING TWELVE EVERY NIGHT.' (6/24/77)

While the substitution of a WH-cleft in these examples would not produce any obvious 'stellar ungrammaticality', it would make them less coherent—since it would, with no justification, presume that the hearer was thinking that one is playing something (33a, in the context of a job description), that a particular woman existed and that she had something (33b, in the context of a description of a dog-bite), and that something kept K.M. from eating the fudge (33c). This last example is an excellent illustration of the difference between GIVEN, i.e. (assumed to be) in the hearer's consciousness, and KNOWN: if questioned, E.P. would have answered correctly that something had kept K.M. from eating the fudge, since it was still there, but this was certainly not on E.P.'s mind. (I am not, however, claiming that *it*-clefts always present known information in the traditional sense.)

[6] It is possible that I am misrepresenting Halliday, as his criteria are not very clear to me. For example, two pages before defining given/new, he states (206): 'Anaphoric items are inherently "given" in the sense that their interpretation depends on identification within the preceding text.' Does this mean that an anaphoric item in a sentence with an unmarked focus is to be independently marked as given? In any event, it does seem that Halliday's notion of givenness does not depend directly on what the speaker takes to be in the hearer's consciousness.

The fact that *it*-clefts make no assumptions about the hearer's thoughts is perhaps more clearly illustrated in cases where the focus is extracted from an idiom:

(34) a. 'It's sort of an arbitrary line that you're drawing.' (D.C., 1/77)
*What you're drawing is sort of an arbitrary line.
b. 'It's obvious that I'm a woman and enjoy being a woman. I'm not overly provocative, either. IT'S THE THIN, GOOD-NIGGER LINE THAT I HAVE TO TOE.' (Writer/producer in Terkel, 105)
*What I have to toe is the thin, good-nigger line.

Since these discourses have nothing to do with drawing or toeing (whatever that is), the speaker could certainly not assume that those activities were being thought about by the hearer.

Investigating a non-feature of *it*-clefts is necessary for the comparison with WH-clefts, but it fails to shed light on what is interesting about them. I shall now turn to what it is that they do.

3.1. MATTERS OF AGREEMENT. As stated above, WH-clefts distinguish between human and non-human focused NP's, whereas *it*-clefts need not:

(35) a. It was golf/John that killed her.
b. What killed her was golf/*John.
(c. The one that killed her was *golf/John.)[7]

Thus it is conceivable that an *it*-cleft may be used in cases where the differentiation imposed by a WH-cleft would make for awkwardness; the following may be examples of such use:

(36) a. '..."The President lied to the American people when he stated that congressional inaction prevented the enactment of this ... legislation." Waxman said. "IT WAS THE PRESIDENT WHO OBSTRUCTED THE PASSAGE OF THIS CRUCIAL BILL"...' (*PB*, 10/8/76, p. 2)
b. 'It didn't make any difference to him whether it was his grant or the department that paid for it.' (P.J.P., 1/77)

If these had WH-clefts instead, then in 36a we would expect that *what* would be parallel to *congressional inaction*, but *the one who* would be in agreement with the following human focus. In 36b, we have a conjunction of non-human and human NP's—which, separately, would display subtle semantic differences in their co-occurrence with *pay*: in *His grant paid for it* or *What paid for it was his grant*, the subject is instrumental, but in *The department paid for it* or *The one that paid for it was the department*, the subject is agentive. These differences, irrelevant to the utterer of 36b, would have made a WH-cleft well-nigh impossible, but they are conveniently ignored in the *it*-cleft. (See fn. 3, above.)

Agreement in number may also be too cumbersome in a WH-cleft, but it presents no problems for the *it*-cleft:

[7] The situation is actually more complex; see fn. 3, above. In addition, as Dwight Bolinger has pointed out to me, *it*-clefts may indeed distinguish between human and non-human foci:
(a) It was golf which/*who killed her.
(b) It was John *which/who killed her.

Austin Hall

(37) a. 'It's not just one person that's hurt. It's usually four or five.' (Heavy equipment operator in Terkel, 51)
b. *The one/ones that is/are hurt isn't/aren't just one person ...

No pattern of agreement is available for 37b which would produce an acceptable sentence.

Although the contrasts presented in this section are strong, I am actually rather hesitant in counting agreement as a definite factor, because I have found no instance of a speaker switching from a cumbersome or impossible WH-cleft to an *it*-cleft. The ideas are offered here solely in the hope that further research, perhaps of an experimental nature, can decide the issue.

3.2. *IT*-CLEFTS IN DISCOURSE. The *it*-cleft construction seems to have quite a few different functions in discourse, and what follows is far from a definitive treatment.

3.21. STRESSED FOCUS *IT*-CLEFTS. Considering the simplest case, in which only the focus has strong stress, we find the *it*-cleft used when the focus represents new, often contrastive information—as in the case of the WH-cleft. The difference lies in the communicative value of the presupposition. In the WH-cleft, as argued above, the WH-clause represents what the speaker is assuming the hearer is or can be thinking about; as such, it may be seen as the THEME of that particular stretch of discourse. In the stressed-focus *it*-cleft, the *that*-clause represents known or old information, which is not marked as assumed to be in the hearer's consciousness and which is not the theme. The contrast is close to that made by Halliday (212) between GIVEN (what I am calling known or old) and THEME:

> 'The difference can perhaps be best summarized by the observation that, while "given" means "what you were talking about" (or "what I was talking about before"), "theme" means "what I am talking about" (or "what I am talking about now"); and, as any student of rhetoric knows, the two do not necessarily coincide.'

Of course, Halliday is speaking of theme in general; the situation with WH-clefts—where the theme is not only what the speaker is talking about, but also what s/he expects the hearer to be thinking about—is special to them. In any event, the *that*-clause of this class of *it*-clefts qualifies as Halliday's given, but not as his theme. Consider the following:

(38) a. '... So I learned to sew books. They're really good books. IT'S JUST THE COVERS THAT ARE ROTTEN.' (Bookbinder in Terkel, 409)
b. ... So I learned to sew books. They're really good books. WHAT'S ROTTEN IS JUST THE COVERS.

In 38a, it is known or knowable, from the fact that the speaker sews books, that something is 'rotten' or amiss with them. However, the theme of this discourse, i.e. what it is about, is 'books', not 'what is rotten'; thus the *it*-cleft is ideally suited to the task.

The stressed-focus *it*-cleft is also ideally suited for illustrating the Prague School notion of COMMUNICATIVE DYNAMISM (CD): 'By the degree of CD carried by a sentence element we understand the extent to which the sentence element contributes to the development of the communication, to which "it pushes the communication forward", as it were' (Firbas 1964:270).[8] *That*-clauses in this class of *it*-clefts can be seen to have a very low degree of CD—so low, in fact, that they are very frequently deleted or are simply missing, yielding a structure that is not obviously formally identifiable as an *it*-cleft but that can be nothing else. Consider 37a above and the following:[9]

(39) a. 'Who made this mold? WAS IT THE TEACHERS? WAS IT THE MEDICINE MAN?' (*Columns*, p. 16)
b. Haldeman: '... I think he will do just everything he can not to hurt the President.'
Nixon: 'Yeah, that has to be true of everybody because IT ISN'T THE MAN, IT'S THE OFFICE.' (*PT*, 267)
c. [general description of routine]
'IT WAS USUALLY TWO TRICKS A NIGHT. That was usually a hundred, a hundred and a quarter.' (Hooker in Terkel, 97)[10]

Stressed-focus *it*-clefts are rather strange in one respect. As has often been noted, in English at least, old information generally precedes new information. WH-clefts, as seen above, conform to this very general pattern. However, consider *it*-clefts like the following:

(40) a. C.B.: 'So who's Barbara?'
B.S.: 'Let me put it this way. When you last saw me with anyone, it was Barbara I was with.' (11/76)
b. '... I was with Barbara.
c. '... the one I was with was Barbara.

The *it*-cleft in 40a extracts the new information, *Barbara*, from final position and places it BEFORE the old information, ∅ *I was with*. Surprisingly, the non-clefted form, 40b, which has everything in the 'right' order, is less natural. I think that the problem with 40b is that the delineation between what is old and what is new is unclear and misleading: one tends to interpret the new information as being *with Barbara* or *was with Barbara*, yielding a rather bizarre reading in which *Barbara* does not identify the *anyone* of the preceding clause. Thus, though the *it*-cleft presents information (old vs. new) in an aberrant order, it clearly marks which is which. The comparison between 40a and 40c, the closest thing possible to a WH-cleft, is more complex. I find 40c sayable, though somewhat less natural than 40a; and I think that once again we have a problem of theme: this has been a discourse about 'Barbara', and 40c would give the impression that the speaker is changing the theme to 'the one I was with'—who, at that moment, could easily be assumed to be 'non-Barbara'. It is perhaps because of such situations that the focus of a

[8] However, Firbas then defines theme as the element having the lowest degree of CD. That might give good results with stressed-focus *it*-clefts—in which *it*, being a dummy, certainly has the lowest degree of CD and would mark the sentence as themeless. However, we shall see below that there is another type of *it*-cleft where the focus may possibly be the theme, but where *it* would necessarily still have the lowest degree of CD.

[9] In Bolinger 1972a,b, such sentences are unreservedly counted as *it*-clefts.

[10] Parenthetically, it is worth noting that, if such sentences are to be derived from full *it*-clefts, it would be increasingly difficult, going from 37a to 39a to 39c, to claim that the deletion rule involved was one based on syntactic identity.

AUSTIN HAIK

WH-cleft must be distinctly less active in the hearer's (assumed) consciousness than the information in the WH-clause: possibly that which is most prominent in the consciousness is (usually?) the theme; and in WH-clefts, the WH-clause is a marked theme.

3.22. INFORMATIVE-PRESUPPOSITION *IT*-CLEFTS. The fact that *it*-clefts may present information as known without making any claims that the hearer is thinking about it (or, in fact, even knows it) presents the speaker with a strong rhetorical temptation: what is to prevent him/her from putting new information into the *that*-clause? Note that such an action on the part of the speaker would be totally in line with the general convention of putting new information last. It turns out that nothing prevents such a use of *it*-clefts; we may distinguish a whole class of sentences in which this occurs, which I shall call INFORMATIVE-PRESUPPOSITION *it*-clefts. With these sentences, not only is the hearer not expected to be thinking about the information in the *that*-clause, but s/he is not expected even to KNOW it. In fact, the whole point of these sentences is to INFORM the hearer of that very information. As an illustration, consider 11 above and discourses like the following:

(41) a. ##'IT WAS JUST ABOUT 50 YEARS AGO THAT HENRY FORD GAVE US THE WEEKEND. On September 25, 1926, in a somewhat shocking move for that time, he decided to establish a 40-hour work week, giving his employees two days off instead of one.' (*PB*, 1/3/76, p. 3L)

b. 'The leaders of the militant homophile movement in America generally have been young people. IT WAS THEY WHO FOUGHT BACK DURING A VIOLENT POLICE RAID ON A GREENWICH VILLAGE BAR IN 1969, AN INCIDENT FROM WHICH MANY GAYS DATE THE BIRTH OF THE MODERN CRUSADE FOR HOMOSEXUAL RIGHTS. (*PG*, p. 16)

c. 'It was also during these centuries that a vast internal migration (mostly by the Galla) from the south northwards took place, a process no less momentous than the Amhara expansion southwards during the last part of the nineteenth century and the beginning of the twentieth century.' (*Challenge*, p. 39)

The discourse in 41a is a newspaper filler. Were the first sentence not clefted, i.e. *Fifty years ago, H.F. gave* ..., it would seem as though the newspaper had just discovered (or were pretending to have discovered) the information in the *that*-clause; the *it*-cleft, in contrast, serves to mark it as a known fact, unknown only to the readership.[11] If we compare 41a to discourse-initial WH-clefts, e.g. 14 above, the difference is striking. In the WH-cleft, what is presupposed on the logico-semantic level is truly presumed on the discourse (pragmatic?) level; but what is presupposed logico-semantically in the informative-presupposition *it*-cleft is NEW information on the discourse level—and therefore, in contrast to both the WH-cleft and the stressed-focus *it*-cleft, has a high degree of CD. The same is true of 41b–c. Such sentences provide strong evidence for distinguishing between these two levels: both are necessary to describe discourse, and each is distinctive.

[11] Although it is known, it is not expected to be known by many people, if we are to believe the newspaper's slogan: 'Nearly everybody reads the *Bulletin*!'

Informative-presupposition *it*-clefts are formally and unambiguously identifiable. First, unlike stressed-focus *it*-clefts, they have normally (vs. weakly) stressed *that*-clauses.[12] Second, they generally have a short and anaphoric focus, which, in my data, is either a (subject) NP or an adverbial, generally of time, place, or reason—Kuno's 'thematic scene-setting' adverbials (1975, 1978). Third, in such sentences, *that/wh-* is not deletable. The sentences in 41 may not be the most convincing evidence for this claim, since 41a,c have a focused adverbial, and 41b has a focused subject. But it is often possible to delete *that* after a focused adverbial, so long as the construction is of the stressed-focus type:

(42) a. [cupping cheeks] 'It's HERE I look like Mina Davis.' (*The last tycoon* [film])

b. 'They don't know how to treat me ... If I had short hair, wore pants, and talked in a gruff voice, that would be more acceptable ... It's WHEN I TRANSCEND THEIR LABELS THEY DON'T QUITE KNOW WHAT TO DO.' (Writer/producer in Terkel, 105–6)

The contrast would be clearer if an object NP were focused in an informative-presupposition *it*-cleft; unfortunately (for my purposes), I have found no such example. It is perhaps significant that all the *it*-clefts I have gathered which clearly have both a focused NP and an informative presupposition—over fifty examples—have a focused initial subject. While I certainly am not claiming that focused objects never occur in this construction, I do consider this an important area for further investigation. In the meantime, let us test the *that*-deletion claim by modifying 41b in such a way as to make it a potential deletion site. *That*/WH- is generally deletable whenever something other than an initial subject has been extracted. Thus we may change 41b to 43b, on the analogy of 43a:

(43) a. I want you to agree with me, no matter who it is that/∅ you think is right.[13]

b. ... It was they who/*∅ everyone knows fought back during ...

(Note that 43b must be read with the stress pattern of an informative-presupposition *it*-cleft, as in 41b. Otherwise, the deletion is predictably acceptable.)

The obvious question that remains is: Why does new information occur in these logically presupposed *that*-clauses? Is it a performance error on the part of the speaker, or is it purposeful and functional? That it is not attributable to error is patently clear: first, these sentences have distinctive, predictable stress patterns; second, they tend to occur in formal, often written, discourse—the type least likely to contain (non-hypercorrective) errors. In fact, they are well-formed and functional. Their function, or at least one of their functions, is TO MARK A PIECE OF INFORMATION AS FACT, known to some people although not yet known to the

[12] The only mention of this class of sentences that I have seen in the literature is by Schmerling (1971:252, fn. 4). She notes their distinctive intonation and describes them as being 'used for rhetorical effect'. She brings them up as a possible counter-example to her thesis that asserted items have heavy stress and presupposed items reduced stress, and she concludes that they are not a counter-example because their *that*-clauses are 'without the attendant presuppositions of the sort described here'. I believe that she is confusing the logico-semantic notions of presupposition/assertion with the discourse notions of given/new information.

[13] For some (probably related) reason, *who* would seem very marginal here.

intended hearer. Thus they are frequent in historical narrative, or wherever the speaker wishes to indicate that s/he does not wish to take personal responsibility for the truth or originality of the statement being made. In a sense, then, they are similar to hedges—e.g. *It seems that, I think that, sorta*—since both have the effect of reducing the speaker's responsibility. Hedges do this by weakening the statement, by making it into a 'guess' or conjecture; *it*-clefts do it by strengthening the statement, by presenting it as an already known fact. Thus the *it*-cleft is particularly well-suited to persuasive discourse, as in 11 and 41c above, and in the following:

(44) a. 'It is through these conquests that THE PEASANTRY BECAME ABSORBED INTO A SINGLE FORM OF DEPENDENT LORD–TENANT RELATIONSHIP.' (*Challenge*, p. 19)

b. 'It is chiefly for the sake of this super-sense, for the sake of complete consistency, that IT IS NECESSARY FOR TOTALITARIANISM TO DESTROY EVERY TRACE OF WHAT WE COMMONLY CALL HUMAN DIGNITY.' (Arendt, 458)

c. 'It is for this reason that HALLE'S ARGUMENT AGAINST AUTONOMOUS PHONEMICS (HALLE, 1959) IS OF SUCH IMPORTANCE: it demonstrates that ...' (Chomsky 1972:127)

Of course, one man's opinion/evaluation is another man's fact, and users of this construction may well be in good faith. It is, therefore, very difficult to draw a line between a propagandistic use of *it*-clefts and a factual-narrative use:

(45) 'It was in this year that YEKUNO AMLAK, A LOCAL CHIEFTAIN IN THE AMBA-SEL AREA, ACCEDED TO THE SO-CALLED SOLOMONIC THRONE.' (*Challenge*, p. 38)

Such *it*-clefts function like traditional footnotes in that they seem to say: 'Don't argue with me—I didn't invent this—and I'm aware that I didn't invent this.'

Another, far more striking analogy that comes to mind is Whorf's description of Hopi (1938:114):

> 'Hopi verbs have three assertions: REPORTIVE (zero form), EXPECTIVE (suffix *-ni*), NOMIC (suffix *-y^{w}i*). These translate, more or less, the English tenses. But they do not refer to time or duration. They distinguish three different kinds of information. Assertion ... is a classification that refers to one of three distinct realms of validity. The reportive is simply a reporting statement, telling of the historical actuality of a certain situation ... The expective declares an expectation or anticipation of a situation ... Since the assertions have nothing to do with time as such, an expectant attitude may be projected into an account of past events ... THE NOMIC DOES NOT DECLARE ANY PARTICULAR SITUATION, BUT OFFERS THE STATEMENT AS GENERAL TRUTH ...' [emphasis supplied]

Whorf further explains that the reportive may be translated by the present or past, depending on whether the speaker is seeing or has seen what is being reported. Thus, using the reportive (= unmarked) form, the speaker is assuming full responsibility for the truth of what s/he is saying. The expective imposes far less responsibility, since what is being asserted is stated as an expectation or conjecture. Likewise the nomic weakens the speaker's responsibility, this time by strengthening the assertion to the point where its truth is certain and in no way depends on the speaker's experience or evaluative talents.

Whorf repeatedly contrasts Hopi's assertion category with the tenses of English; he considers the contrast an example of the diversity of languages, an important theme in his time. However, the contrast may be less significant than it at first appears: in both English and Hopi, speakers can indicate how much responsibility they wish to take for the validity of their statements. In Hopi, they do it obligatorily by affixation; in English, they do it optionally in a variety of ways—hedges, 'future auxiliary markers',[14] *it*-clefts etc. The contrast is actually best described in Whorfian terms; i.e., both English and Hopi have categories of responsibility-marking. But in English it is covert, whereas in Hopi it is overt (Whorf 1945:88–9):

> 'AN OVERT CATEGORY is a category having a formal mark which is present ... in every sentence containing a member of the category ... A COVERT CATEGORY is marked, whether morphemically or by sentence-pattern, only in certain types of sentences and not in every sentence in which a word or element belonging to the category occurs.'

The fact that, in both English and Hopi, speakers can indicate how much responsibility they wish to take for what they are saying should not be seen as a startling coincidence; it is not unreasonable to hypothesize that this is a possibility in every language—though, of course, the particular mechanisms will differ. Thus Fought (1976:234–5) points out that, in Chorti, a Mayan language of Guatemala, a particle, *txe*, has just such a function:

> 'The form *txe* 'he said' or 'it is said' occurs (as a narrative particle) at the beginning of an event cluster; it can be said to invoke the authority of an anonymous past narrator (the person who told the story to the person who is telling it to you), or just 'people'. It was that past narrator who said it this way; so I too shall say it ... It typically comes near the beginning and end of a tale, presumably as a means of claiming for the narrative the authority of the old bearers of tradition, and then disclaiming personal responsibility when retelling the story.'

Isidro Gonzáles M., Fought's Chorti storyteller, and William L. Shirer, the author of *The rise and fall of the Third Reich*, have far more in common, both in terms of their societal roles and professional output and (therefore) in their style of discourse, than might at first be suspected.

Thus all informative-presupposition *it*-clefts have the general function of presenting statements as facts; but some also have more specific, related sub-functions:

(46) a. ##'IT WAS 10 YEARS AGO THIS MONTH THAT YOUNG IRWIN VAMPLEW WAS BOPPED ON THE HEAD BY A NIGHTSTICK WHILE SMASHING WINDOWS IN BERKELEY IN ORDER TO END THE WAR IN VIETNAM. So you can imagine the elation of his parents when he finally emerged this week from his decade-long coma.

'His first words, naturally, were: "Down with the Establishment!"' (*PI*, 2/6/77, p. 7D)

b. ##'IT WAS JUST A YEAR AGO THAT THE CITY'S MAJOR BANKS LAUNCHED ... THE PHILADELPHIA MORTGAGE PLAN ... TO FOSTER INVESTMENT IN OLDER NEIGHBORHOODS.

'Today, PMP is still in business and going strong. The thirteen ... banks ... have 1,037 mortgages in force for a total of $10.7 million.' (*PI*, 10/22/76, p. 1K)

[14] Compare the different effects of each of the following:

(a) I {will leave, am going to leave, am leaving, am to leave ...} tomorrow.

(b) {to houseguest} How long will you {stay, be staying}?

The *it*-clefts in 46 have the (sub-)function of bringing the reader up to date so that s/he can appreciate the actual 'news', in both cases set off as the second (and final) paragraph. A more subtle use of such an *it*-cleft, where it is used ironically, is:

(47) ##'IT IS THROUGH THE WRITINGS OF BASIL BERNSTEIN THAT MANY SOCIAL SCIENTISTS HAVE BECOME AWARE OF THE SCIENTIFIC POTENTIAL OF SOCIOLINGUISTICS ...
'Yet their very popularity has often deformed Bernstein's arguments; ... he has been made to say that lower class children are linguistically "deprived" ... In fact, Bernstein's views are much more complex than that. First, ...' (Giglioli 1972:155)

Note that, in both 46 and 47, the information in the *it*-clefts, though discourse-initial, is background material—subordinate in importance (in terms of what the discourse is about) to what follows.

A subordinating sub-function often plays a role in other types of informative-presupposition *it*-clefts as well, in that the *it*-cleft may suggest that one state of affairs is closely tied to another, often implicating a cause-and-effect relationship:

(48) a. 'Here ... were the ideas which Hitler was later to use ... His originality lay in his being the only politician of the Right to apply them to the German scene after the First World War. IT WAS THEN THAT THE NAZI MOVEMENT, ALONE AMONG THE NATIONALIST AND CONSERVATIVE PARTIES, GAINED A GREAT MASS FOLLOWING AND, HAVING ACHIEVED THIS, WON OVER THE SUPPORT OF THE ARMY, THE PRESIDENT OF THE REPUBLIC AND ... BIG BUSINESS—THREE "LONG-ESTABLISHED INSTITUTIONS" OF GREAT POWER. The lessons learned in Vienna proved very useful indeed.' (Shirer, 24–5)
b. 'But why is the topic so important? Apparently, IT IS THE TOPIC THAT ENABLES THE LISTENER TO COMPUTE THE INTENDED ANTECEDENTS OF EACH SENTENCE IN THE PARAGRAPH.' (Clark & Haviland, 49)

If the third sentence of 48a read *Then, the Nazi movement* ..., it would tend to suggest a separate event, and we would lose the notion that it was all H's doing—a notion conveyed very strongly by the *it*-cleft's subordinating effect, and underlined (though still not asserted) by the last sentence. In 48b, the *it*-cleft obviates the need for *because*, without which the unclefted version would be very awkward.

A third sub-function of informative-presupposition *it*-clefts is that of indicating deference/politeness—a surprise, perhaps, to those who think they are the exclusive property of pompous rhetoricians. Actually, it is not so surprising. Consider the following:

(49) a. 'However, it turns out that there is rather interesting independent evidence for this rule, and IT IS TO THAT EVIDENCE THAT WE MUST NOW TURN.' (Akmajian & Wasow 1975:221)
b. 'It's with great honor and pleasure that I announce Hilary Putnam.' (grad. stud., U. of P., 4/1/77)

In sentences like 49a, which are very frequent in the data, the authors remove themselves as subject (and as potential theme); at the same time, they seem to present their act of turning as an inescapable, external fact—not as their own choice. (Note, in this regard, the occurrence of the analogously-functioning hedge *must*.) In 49b, an equally common type, we again find the speaker playing down his own role in the event, subordinating it and himself in the situation as in the syntactic structure. Note here that the *that*-clause contains a performative sentence. I leave open the question of what that does to the notion that *it*-cleft *that*-clauses are logically presupposed—i.e. true if the sentence is meaningful—in a theory in which performative sentences are considered to be neither true nor false, but felicitous or infelicitous. In any event, it deals one more blow to the notion that *it*-clefts and WH-clefts are equivalent, and that *it*-cleft *that*-clauses represent old information in the discourse.

4. GIVEN VS. KNOWN. These terms are generally used interchangeably in the literature; at the risk of adding to an already large amount of confusion, I should like to propose that they be used to denote separate things:

(50) a. GIVEN INFORMATION: Information which the coöperative speaker may assume is appropriately in the hearer's consciousness.
b. KNOWN INFORMATION: Information which the speaker represents as being factual and as already known to certain persons (often not including the hearer).

As defined here, GIVEN and KNOWN are very different. Given information is relevant to WH-clefting, as argued above, and also to such phenomena as pronominalization and probably all types of so-called 'deletion under identity'. Right dislocation is perhaps a correction of the speaker's initial assumptions of what can be taken as given. In general, if the speaker is uncoöperative about given information, the resulting utterance ranges from clumsy, to incomprehensible, to what the layman (and perhaps also the psychiatrist) would describe as schizophrenic. The speaker is seen as being out of touch with the reality of the speech situation.

Known information, however, is less related to what one thinks of as 'grammar', and less immediately relevant to speaker–hearer interaction. It is rather a choice on the part of the speaker of a particular validity-level that s/he wishes to ascribe to the utterance. Markers of this choice (corresponding to Whorf's nomic) include the *it*-cleft and particular word choices (*the fact that*, factive predicates etc.).[15] Looseness in designating information as known (i.e. over-designating) tends to yield an effect of pomposity or manipulativeness, rather than incomprehensibility. Paradoxically, known information may also signal deference, for the same reason that hedges do: both lessen the speaker's responsibility for the statement made; they reduce the speaker's assertiveness, claims to originality—and, therefore, role.

The distinction between given information, as carried by WH-clefts, and known information, as carried by *it*-clefts, may account for some of the results obtained in an experiment reported in Hornby 1974. Subjects were presented with a series of 60 pre-recorded sentences, each followed by a picture, and were asked to say

[15] This is directly in line with Bolinger's statement (1977:68): 'A factive verb implies the factuality of its complement in the mind of the hearer, not the shared knowledge of it between speaker and hearer.' I thank him for sending me a copy of the chapter on *it*, which is strikingly compatible with many of the ideas expressed above.

Austin date

whether each picture accurately depicted the preceding sentence. Of the 60 sentences, 20 were simple, e.g. *The girl is petting the cat*; 20 were *it*-clefts, e.g. *It is the girl that is petting the cat*; and 20 were akin to WH-clefts, e.g. *The one that is petting the cat is the girl*. Of each set of 20 sentences, 10 were active, 10 passive. Of each set of 10 sentences, 5 were followed by pictures which misrepresented their presupposition, and 5 by pictures which depicted them accurately. The results in Table 2 below show the percentage of times that the subjects failed to note a misrepresentation, i.e. said the picture was an accurate depiction when it was not.

	ACTIVE	PASSIVE	TOTAL
SIMPLE	15.6%	40.6%	28.2%
IT-CLEFT	30.6%	48.0%	39.4%
WH-CLEFT	44.6%	55.2%	50.2%

TABLE 2. Subjects' failure to note misrepresentation (Hornby 1974).

Immediately striking and deserving of further study is the difference between active and passive. Relevant to the issue at hand, however, are the differences displayed vertically, and I shall limit the rest of my discussion to the active sentences. I assume that the subjects were coöperative, and attempted to the best of their ability to react to an experiment-sentence as though it had occurred naturally. If this is so, then a simple sentence, being merely the speaker's assertion, would elicit the most verification; and in fact discrepancies were noted in 85% of the cases. An *it*-cleft (these were all of the stressed-focus type) marks the *that*-clause information as known, which would tempt the subject to be less careful about verifying it; in fact, discrepancies were missed twice as often. In the WH-cleft, the information in the WH-clause is marked as given, i.e. assumed to be in the hearer's consciousness; and the coöperative subjects, acting as though that were the case, did not bother to verify in nearly half the cases.[16]

5. CONCLUSION. I have tried to show that, far from being interchangeable, *it*-clefts and WH-clefts have highly specialized distributions and functions: WH-clefts mark the information in the WH-clause as assumed/assumable to be in the hearer's consciousness, or GIVEN. What can be given depends on the linguistic and non-linguistic context and on the speech situation. *It*-clefts, however, are far more heterogeneous, but all types mark the information in the *that*-clause as a known fact—or simply KNOWN. The two types of *it*-clefts that have been discussed here are (a) the stressed-focus *it*-cleft, in which the focus represents new information, and the *that*-clause represents information which is often, though not always, known from the context; and (b) the informative-presupposition *it*-cleft, in which the focus usually contains an anaphoric item, and the *that*-clause contains the 'message'—but marked as a known fact, not as the speaker's assessment. The latter is typically used in historical narrative and at other times when self-effacement is sought by the speaker, e.g. for politeness. This type of *it*-cleft may also have a subordinating function, sometimes to indicate 'backgroundness', sometimes to implicate cause and effect. The functions mentioned here are not intended as an exhaustive list; in

[16] While the results are clear, I still have doubts about the appropriateness of the test situation—using sentences in isolation—as an effective means of discovering how speakers verify statements in natural discourse. Perhaps a less artificial experiment can be devised.

particular, I have barely touched on the notion of theme, which is obviously closely related. Likewise, the types of *it*-clefts listed here are not meant to be exhaustive; one other type is the 'proverbial' *it*-cleft, for which it is difficult to say that the *that*-clause is presupposed:

(51) a. It's a wise child that knows its own father.
b. It's a fortunate man that will find the fountain of youth.[17]

Doubtless there are other types as well.

One different but closely related construction that should be studied is the inverted WH-cleft; e.g.,

(52) a. That is what I want.
b. John is who I saw.
c. Paris is where he met his wife.

I would predict that the WH-clause in such sentences is marked not as GIVEN but as KNOWN; a comparison with both WH-clefts and *it*-clefts is very much in order. The one final point I wish to stress is the importance of studying these and all other sentence constructions IN DISCOURSE, since it is only there that their communicative functions can be observed.

REFERENCES

AKMAJIAN, ADRIAN. 1970. On deriving cleft sentences from pseudo-cleft sentences. LI 1.149–68.

——, and FRANK HENY. 1975. An introduction to the principles of transformational syntax. Cambridge, MA: MIT Press.

——, and THOMAS WASOW. 1975. The constituent structure of VP and AUX and the position of the verb BE. Linguistic Analysis 1.205–46.

ARENDT, HANNAH. 1951. The origins of totalitarianism. New York: Harcourt Brace.

BALL, CATHERINE N. 1977. Th-clefts. Paper read at the University of Pennsylvania Linguistics Colloquium, Philadelphia.

BOLINGER, DWIGHT. 1972a. A look at equations and cleft sentences. Studies for Einar Haugen, ed. by E. Firchow, 96–114. The Hague: Mouton.

——. 1972b. That's that. The Hague: Mouton.

——. 1977. Meaning and form. London: Longman.

CHAFE, WALLACE. 1974. Language and consciousness. Lg. 50.111–33.

——. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. Subject and topic, ed. by Charles Li, 25–56. New York: Academic Press.

CHALLENGE. Journal of the World-Wide Union of Ethiopian students. Vol. X, No. 1, February 1970.

CHOMSKY, NOAM. 1970. Deep structure, surface structure, and semantic interpretation. Studies in general and oriental linguistics presented to Shiro Hattori, ed. by R. Jakobson & S. Kawamoto, 52–91. Tokyo: TEC. [Reprinted in his Studies on semantics in generative grammar, 62–119. The Hague: Mouton, 1972.]

——. 1972. Some empirical issues in the theory of transformational grammar. In his Studies on semantics in generative grammar, 120–202. The Hague: Mouton.

CLARK, HERBERT, and SUSAN E. HAVILAND. 1977. Comprehension and the given-new contrast. Discourse production and comprehension, ed. by Roy Freedle, 1–40. Hillside, NJ: Erlbaum.

COLUMNS. Univ. of Pennsylvania magazine. March 1977.

EMONDS, JOSEPH. 1970. Root and structure-preserving transformations. MIT dissertation.

[17] These were brought to my attention by Catherine N. Ball (see Ball 1977 for a discussion).

FIRBAS, JAN. 1964. On defining the theme in functional sentence analysis. Travaux Linguistiques de Prague 1.267–80.
FOUGHT, JOHN G. 1976. Time structuring in Chorti (Mayan) narratives. Mayan linguistics I, ed. by Marlys McClaran, 228–42. Los Angeles: American Indian Studies Center, UCLA.
GIGLIOLI, PAOLO (ed.) 1972. Language and social context. Harmondsworth: Penguin.
GREEN, GEORGIA. 1976. Main clause phenomena in subordinate clauses. Lg. 52.382–97.
GRICE, H. P. 1975. Logic and conversation. Syntax and semantics III: Speech acts, ed. by Peter Cole & Jerry Morgan 41–58. New York: Academic Press.
GUNDEL, JEANETTE K. 1977. Where do cleft sentences come from? Lg. 53.343–59.
HALLIDAY, M. A. K. 1967. Notes on transitivity and theme in English, Part 2. JL 3.199–244.
HANKAMER, JORGE. 1974. On the non-cyclic nature of WH-clefting. Proceedings of the 10th Regional Meeting, Chicago Linguistic Society, 221–33.
HAVILAND, SUSAN E., and HERBERT CLARK. 1974. What's new? Acquiring new information as a process in comprehension. JVLVB 13.512–38.
HIGGINS, F. R. 1971. The pseudo-cleft construction in English. MIT dissertation.
HOOPER, JOAN, and SANDRA THOMPSON. 1973. On the applicability of root transformations. LI 4.465–97.
HORNBY, PETER A. 1974. Surface structure and presupposition. JVLVB 13.530–38.
JE. Jewish exponent (Philadelphia weekly newspaper).
KEENAN, EDWARD. 1971. Two kinds of presupposition in natural language. Studies in linguistic semantics, ed. by Charles Fillmore & D. T. Langendoen, 45–54. New York: Holt.
KLIMA, EDWARD, and URSULA BELLUGI-KLIMA. 1969. Syntactic regularities in the speech of children. Modern studies in English, ed. by D. Reibel & S. Schane, 448–66. Englewood Cliffs, NJ: Prentice-Hall.
KUNO, SUSUMU. 1975. Conditions for verb phrase deletion. FL 13.161–75.
——. 1978. Two topics on discourse principles. To appear in Bulletin of the ICU Summer Institute in Linguistics, 2.
LASNIK, HOWARD. 1976. Remarks on coreference. LA 2.1–22.
MERRITT, MARILYN. 1976. Resources for saying in service encounters. U. of Pennsylvania dissertation.
PB. Philadelphia/Sunday bulletin (daily newspaper).
PG. Pennsylvania Gazette (U. of Penn. alumni magazine). February 1977.
PI. Philadelphia/Sunday inquirer (daily newspaper).
PINKHAM, JESSIE, and JORGE HANKAMER. 1975. Deep and shallow clefts. Proceedings of the 11th Regional Meeting, Chicago Linguistic Society, 429–50.
PRINCE, ELLEN F. 1976. The syntax and semantics of NEG-raising, with evidence from French. Lg. 52.404–26.
PT. The presidential transcripts. Prepared by the Washington Post. New York: Dell, 1974.
QUIRK, RANDOLPH, and SIDNEY GREENBAUM. 1973. A concise grammar of contemporary English. New York: Harcourt Brace Jovanovich.
SHIRER, WILLIAM L. 1960. The rise and fall of the Third Reich. New York: Simon & Schuster.
SCHMERLING, SUSAN. 1971. Presupposition and the notion of normal stress. Proceedings of the 7th Regional Meeting, Chicago Linguistic Society, 242–53.
TERKEL, STUDS. 1974. Working. New York: Pantheon. [Reprinted, New York: Avon, 1975.]
TODAY. Magazine of the Sunday inquirer.
WHORF, BENJAMIN. 1938. Some verbal categories of Hopi. Lg. 14.275–86. [Reprinted in his Language, thought, and reality, 112–24. Cambridge, MA: MIT Press, 1956.]
——. 1945. Grammatical categories. Lg. 21.1–11. [Reprinted in his Language, thought, and reality, 87–101.]

[Received 3 November 1977.]